Institute for Scientific Computing Research

FY 1998 Seminar Series

(in reverse chronological order)

September 29,1998

Hybrid Approaches for Solution of Large-Scale Systems of Nonlinear Equations

Michael Pernice

University of Utah
pernice@sneffels.chpc.utah.edu

Abstract

There has been considerable discussion of the relative merits of multigrid and Newton-based methods for solving large-scale systems of nonlinear equations. The full approximation scheme (FAS) promises cheap iterations and high rates of convergence, but this is difficult to achieve without the correct combination of interlevel transfers and smoothing strategies. While superlinear rates of convergence are attractive, Newton-based iterations are more expensive (both in terms of operations and storage) and even with globalization strategies performance can be sensitive to the choice of an initial approximation. Recent advances in the development of inexact Newton methods have made this approach more competetive, but a good preconditioner is still necessary to achieve satisfactory performance.

This debate is largely misdirected. Since both approaches require selection of components whose effectiveness is highly problem-dependent, it is unlikely that one strategy will emerge as the method of choice for a broad class of applications that involve complex multiphysics and multiscale phenomena. In fact, the two approaches have complementary strengths and weaknesses that can be exploited to compose efficient and robust strategies for solving nonlinear systems of equations. The expense of a Newton-Krylov method can be mitigated by using it as the coarse grid solver for FAS, which will often provide a good initial approximation. An improved coarse grid solver can also improve the robustness of a nonlinear multigrid scheme. Linear multigrid methods can be used as preconditioners for a Newton-Krylov scheme, whether it is used as a coarse grid solver or as a standalone solver. As a standalone solver, Newton-Krylov schemes with multigrid preconditioning can make the choice of a smoother less critical. Further, no explicit linearization is needed to make these strategies effective.

Even within this framework, numerous questions still need to be addressed. What is the best multigrid cycling strategy to use as a preconditioner? How accurately should the FAS coarse grid problem be solved? Where in the grid hierarchy is the Newton-Krylov method most effective? What software components are needed, and how should they be organized, to facilitate exploration of these issues? Few analytic results are available to point to the most effective strategy, but numerical experiments help to identify fruitful avenues to investigate.

University of California
Lawrence Livermore National Laboratory

ISCR Institute for Scientific Computing Research

September 28, 1998

Parallel PIC Modeling of Semiclassical Quantum Models

Dean Dauger

University of California, Los Angeles
dauger@physics.ucla.edu

Abstract

A new simulation model based on the semiclassical approach has been constructed to investigate the dynamics of quantum mechanical systems. The simulation code, which is under development, uses classical particle computations to approximate Feynman path integrals that evolve quantum mechanical wavefunctions forward in time. Interactions are treated semiclassically. The algorithm for the evolution of the classical particle trajectories is based on a plasma particle-in-cell code that is designed for parallel computers and is scalable to many processors. This will facilitate the modeling of many particle (100–1000) quantum systems. Progress will be reported on one-dimensional single and multiparticle systems and comparisons with physical situations with known analytic solutions will also be discussed.

CASC Center for Applied Scientific Computing

September 23, 1998

Fermion Monte Carlo

Mal Kalos

Cornell University
Theory & Simulation Science & Engineering Center

Abstract

We review the fundamental challenge of fermion Monte Carlo for continuous systems, the "sign problem," and some of the proposals that have been made for its solution, including some approximate schemes and others whose computing requirements grow exponentially. The issue is to find not the fundamental eigenmode of the Schrödinger equation in many dimensions, but one that has a special inversion of sign—antisymmetry in the exchange of pairs of coordinates. Thus the function sought is not everywhere positive.

We also describe a class of methods that depend upon the use of correlated dynamics for ensembles of walkers that carry opposite signs. We discuss the algorithmic symmetry between such walkers that must be broken to create a method that is both exact and as effective as for treating the fundamental symmetric solution. We explain the concept of marginally correct dynamics. Stable overlaps with an antisymmetric trial function given by such dynamics correspond to the lowest antisymmetric mode. Many-body harmonic oscillator problems are particularly tractable: Their stochastic dynamics permits the use of regular geometric structures for the ensembles, structures that are stable when appropriate correlations are introduced, and that avoid the decay of signal-to-noise that is a normal characteristic of the sign problem. Finally, we outline a new generalization of the method for arbitrary potentials and describe the success in treating few-electron systems, free fermions and small systems of ^{3}H.

University of California
Lawrence Livermore National Laboratory

September 17, 1998

Developing Effective Communication Abstractions for Clusters of SMPs

Steven Lumetta

University of Illinois at Urbana-Champaign
steve@crhc.uiuc.edu

Abstract

Clusters of SMPs, or Clumps, present a two-level hierarchy for interprocess communication. Effectively addressing this hierarchy requires a substantial effort. Clearly, many applications will require that a programmer recognize the hierarchy to some extent to obtain optimal performance, but an effective abstraction can relieve the programmer of the bulk of the workload.

In the first part of my talk, I present a lightweight message-passing layer that abstracts communication into a single interface, transparently directing communication traffic through the appropriate medium. I describe the design, implementation, and performance of this multi-protocol layer on a cluster of Sun Enterprise 5000 servers connected by Myrinet. Applications within an SMP benefit from the use of fast communication, but resource contention can dominate performance on a Clump. In the second half of my talk, I present models of application performance and argue that realizing the full potential of the Clumps architecture may require that we abandon the global phase structuring that has proven so effective in developing parallel applications.

September 11, 1998

PLUM: Parallel Load Balancing for Adaptive Unstructured Meshes

Leonid Oliker

NASA Ames Research Center
oliker@riacs.edu

Abstract

Dynamic mesh adaption on unstructured grids is a powerful tool for computing unsteady three-dimensional problems that require grid modifications to efficiently resolve solution features. By locally refining and coarsening the mesh to capture flowfield phenomena of interest, such procedures make standard computational methods more cost effective. Unfortunately, the adaptive solution of unsteady problems causes load imbalance among processors on a parallel machine. An efficient parallel implementation of such methods is extremely difficult to achieve, primarily because of the dynamically changing, nonuniform grid.

This talk will present the development of PLUM, an automatic portable framework for performing adaptive numerical computations in a message-passing environment. Mesh adaption, repartitioning, processor assignment, and remapping are critical components of the framework that must be accomplished rapidly and efficiently so as not to cause a significant overhead to the numerical simulation. PLUM requires that all data be globally redistributed after each mesh adaption to achieve load balance. An algorithm is presented for minimizing this remapping overhead by guaranteeing an optimal processor assignment. Results demonstrate that applying a processor reassignment algorithm to the default mapping of the parallel partitioner can significantly reduce the data redistribution cost.

A data redistribution model is also presented that predicts the remapping cost on the SP2. This model is required to determine whether the gain from a balanced workload distribution offsets the cost of data movement. Portability is examined by comparing performance on a SP2, and Origin2000, and a T3E. A CCNUMA implementation of this work in currently underway. Preliminary shared memory results on the Origin2000 will also be presented.

September 3, 1998

GPFS - a Scaleable Parallel File System for the RS/6000 SP

Roger Haskin

IBM Almaden Research Center

Abstract

General Parallel File System (GPFS) is a scaleable parallel file system for the RS/6000 SP. GPFS allows large numbers of disks attached to multiple storage servers to be configured as a single file system. Files are striped across all disks in a file system for high throughput.

Programs running on multiple nodes in the SP can access data in a GPFS file in parallel. In addition to high-speed parallel file access, GPFS provides fault-tolerance. Working in conjunction with the SP2 Phoenix software, GPFS recovers automatically from disk and node failures. Its robust design makes GPFS appropriate for commercial applications such as large Web servers, data mining, and digital libraries. This talk presents the high-level architecture of GPFS, describes some of its advantages, and discusses plans for future work.

August 31, 1998

Rayleigh-Taylor Mixing: Experiments and Simulations

Yuan-Nan Young
and
Jonathan Dursi

University of Chicago

Abstract

Rayleigh-Taylor mixing, which occurs between unstably stratified fluids, is important in such areas as supernovae type Ia and inertial confinement fusion (ICF). Details of the mixing, however, are still poorly understood. We investigate the detailed spatial structure and size of the mixing zone by analyzing data from previous experiments on the Linear Electric Motor (LEM) using immiscible fluids, and results from numerical simulations with incompressible, Boussinesq fluid (low Atwood numbers) using a 2D pseudospectral code. We then compare differences and similarities between numerical and experimental results.

University of California
Lawrence Livermore National Laboratory

ISCR Institute for Scientific Computing Research

August 25, 1998

X–Ray Diffraction and the Holographic Inverse Problem

Abraham Szoke

LLNL Researcher Emeritus
szoke1@llnl.gov

Abstract

X-ray diffraction on crystals is the most important tool for finding the three- dimensional structure of large molecules—proteins and DNA. There is a close parallel between x-ray diffraction and holography. Utilizing this parallel we have written a computer code that solves structures of large molecules. In this talk I will describe x-ray crystallography, holography, inverse problems and protein structure in one easy lesson.

August 24, 1998

Algorithmic Scalability Issues in Computational Fluid Dynamics

Timothy Barth

NASA Ames Research Center
barth@nas.nasa.gov

Abstract

The first half of the presentation will focus on theoretical techniques for analyzing the performance of iterative methods for the CFD equations. In the second half of the presentation, I will discuss some ongoing work in parallel preconditioning for non-selfadjoint partial differential equations (PDEs).

Despite early success in the design of optimal complexity algorithms for elliptic PDE problems using multigrid (MG) and domain decomposition (DD) techniques, the design of optimal complexity algorithms in computational fluid dynamics (CFD) remains largely an open problem. Even among theoretically optimal elliptic PDE algorithms, computer architecture/hardware dependent characteristics such as memory latency and bandwidth conspire to prevent true parallel scalability from being achieved. The situation is more complex for the CFD equations where action at long distances takes place along characteristic directions.

To understand the performance of iterative methods for solving the discretized CFD equation, it is desirable to obtain energy and condition number properties for the symmetric portion of the discrete operator. In the present work, I consider systems of nonlinear conservation laws with convex extension together with stabilized finite element approximation. A general space-time discontinuous function space is assumed. Using nonlinear techniques together with specialized stabilization techniques, it is possible to derive exact energy properties of these stabilized numerical methods for nonlinear conservation laws. Crucial to the analysis is the precise form of stabilization used in the finite element formulations. Using the analysis, subdomain and interface energies are then readily obtained from the theory. The exact energy balance equation is very revealing. For example, when applied to the nonlinear compressible Navier-Stokes equations, the exact analysis clearly identifies the ill-posed nature of the system energy for recirculating flow at high Reynolds number. The analysis also shows subtle deficiencies in the model problem analysis of Yavneh and Brandt for recirculating flow.

In the second portion of the presentation, I will consider preconditioning methods for convection-dominated fluid flow problems based on a nonoverlapping Schur complement domain decomposition procedure for arbitrary triangulated domains. This is joint work with Tony Chan (UCLA) and W.-P. Tang (University of Waterloo). The triangulation is first partitioned into a number of subdomains and interfaces, which induces a natural 2X2 partitioning of the PDE discretization matrix. We view the Schur complement induced by this partitioning as an algebraically derived coarse space approximation. This avoids the known difficulties associated with the direct formation of an effective coarse discretization for advection dominated equations. By considering various approximations of the block factorization of the 2X2 system, we have developed a family of robust preconditioning techniques. These approximations are introduced to improve both the sequential and parallel efficiency of the method without significantly degrading the quality of the preconditioner. The specific approximations that we have used include ILU-preconditioned GMRES subdomain solves, localized approximation of the interface Schur complement, and limited level-fill ILU interface backsolves. A number of 2D CFD calculations will be presented for both scalar advection-diffusion equations and the Euler equations.

ISCR Institute for Scientific Computing Research

August 21, 1998

The Trials and Tribulations of Scriptable Scientific Software

David Beazley

University of Chicago
beazley@cs.cs.utah.edu

Abstract

Scripting languages such as Perl, Python, and Tcl are becoming an increasingly popular tool for the creation of flexible scientific software. Much of this popularity is due to the fact that scripting languages provide scientists with an interpreted environment that can be used for exploratory problem solving similar to that found in packages such as MATLAB, Mathematica, Maple, and IDL. In addition, scripting languages often simplify software development because they encourage the use of software components and provide a high-level environment for debugging and testing.

In this talk, I describe the use of scripting languages with the SPaSM large-scale molecular dynamics code at Los Alamos National Laboratory. Originally developed for the Connection Machine 5, SPaSM initially proved to be too difficult to use and maintain to be of practical value to scientists. However, the use of scripting languages effectively transformed this application into a highly flexible system that is now being used on a daily basis. Various aspects of the scripting environment will be described including an integrated data analysis and visualization component, remote simulation monitoring over the Internet, and the use of scripting languages on parallel machines such as the Avalon DEC-Alpha Linux cluster at Los Alamos (a price-performance finalist in the 1998 Gordon-Bell Prize).

Currently, SPaSM utilizes Python (www.python.org) for its scripting interface. In addition, the SWIG interface generator (www.swig.org) is used to construct Python interfaces to C/C++ libraries. Much of the talk will focus on the impact of using these tools over a three-year period. In addition, limitations and future challenges will be discussed.

August 19, 1998

The SCIRun Problem Solving Environment

Steve Parker

University of Utah ASCI Alliance Center
sparker@taz.cs.utah.edu

Abstract

Computational steering has been defined as "the capacity to control the execution of long-running, resource-intensive programs."

In the field of computational science, we apply this concept to link visualization with computation and geometric design to interactively explore (steer) a simulation in time and/or space. As knowledge is gained, a scientist can change the input conditions and/or other parameters of the simulation.

Although steering was proposed over a decade ago, it is only gradually becoming a popular paradigm for scientific computing. Computational steering is difficult because it requires in-depth knowledge in a wide range of disciplines from geometric modeling to scientific computing to scientific visualization and graphics. Most scientists do not have the necessary expertise in visualization, and most visualization experts do not perform large-scale scientific simulations. In order to successfully apply computational steering to these iterative design problems, we implement a problem solving environment (PSE), called SCIRun, wherein these various phases of the scientific computing process may be integrated.

Implementation of a computational steering framework requires a successful integration of the many aspects of scientific computing, including geometric modeling, numerical analysis, and scientific visualization. All aspects must be effectively coordinated within an efficient computing environment (which, for large-scale problems, means dealing with the subtleties of various high-performance architectures).

In this talk, I will describe how the architecture of SCIRun addresses these problems.

August 14, 1998

Scaling Heterogeneous Information Access for Wide-Area Environments

Louiqa Raschid

University of Maryland
louiqa@umiacs.umd.edu

and

Michael Franklin

University of Maryland
franklin@cs.umd.edu

Abstract

Research in data integration technology aims to enable seamless access to data stored in a wide variety of repositories. Recently, numerous successful prototype systems have been developed. As such systems begin to be deployed in a wide-area network-based environment, however, they will encounter significant challenges arising from the huge number of disparate, unpredictable, and unreliable repositories.

In order for the deployment of such systems to be successful, data integration technology must overcome many scalability problems, including

- The capabilities and contents of heterogeneous repositories can be very dissimilar, making it difficult to generate working access plans.
- The availability and response time for accessing remote repositories can fluctuate dramatically.
- Little support currently exists for identifying and locating repositories that are relevant to a particular application.

In this talk, we describe techniques that we have developed for addressing these scalability problems, including

- Tools for generating wrappers, specifying source capabilities, and generating alternative access plans.
- Query Scrambling, a reactive approach to query execution that can adapt to changes in the environment.
- WebSemantics, an architecture for publishing, locating, and transparent access to sources via the WWW. We then discuss our on-going research on metadata management for the discovery, querying, and dissemination of data sources on the Internet.

August 12, 1998

Real Gas Effects in Oblique Shock Wave Reflection Experiments and Computations

Harland Glaz

University of Maryland Department of Mathematics
hmg@maria.umd.edu

Abstract

An extensive series of shock wave reflection experiments was performed (1970–1990, approximately) at the University of Toronto Institute for Aerospace Studies (UTIAS) under the direction of Professor I. I. Glass. Test gases included air, CO_2, and SF_6; whole flowfield data is available from interferograms. Since 1975, CFD simulations of these experiments have been widely undertaken for various purposes, and with substantial success.

This talk will focus on a series of "anomalous" experiments in heavy test gases (such as SF6 and isobutane) run at high shock-wave Mach number, very low ambient test gas density and pressure, and at or near 37^\degree wedge angle. Interferograms and computational results using second-order Godunov schemes will be presented and compared. It will be seen that much of the phenomenology is still poorly understood. Computational results attempting to partially resolve the situation via inviscid simulations, but with a nonconvex EOS, are presented, including new results by Bei Wang ('98 Ph.D. dissertation, UMD). Other, more speculative, ideas will be discussed as well.

ISCR Institute for Scientific Computing Research

August 7, 1998

Experiment Management Support for Parallel Performance Tuning

Karen L. Karavanic

University of Wisconsin, Madison
karavan@cs.wisc.edu

Abstract

Existing performance tools focus on gathering and reporting information about a single execution of a program. However, the development of a high-performance parallel system or application is an evolutionary process: Both the code and the environment go through many changes during a program's lifetime. At each change, a key question for developers is, how and how much did the performance change?

My research reframes performance tuning as a specialized instance of scientific experimentation, and develops methods for storing, viewing, and using performance data that span a variety of executions, program versions, and environments. In this talk I shall report some early results with a performance tuning study and with a scientific application run in changing environments. I shall also describe work in progress that investigates the use of historical performance data to improve automated performance diagnosis.

July 28,1998

Preconditioning for the Steady-State Navier–Stokes Equations with Low Viscosity

Howard Elman

University of Maryland
elman@cs.umd.edu

Abstract

We present a preconditioner for the linearized Navier–Stokes equations that is effective when either the discretization mesh size or the viscosity approaches zero. For constant coefficient problems with periodic boundary conditions, we show that the preconditioning yields a system with a single eigenvalue equal to one, so that performance is independent of both viscosity and mesh size.

For other boundary conditions, we demonstrate empirically that convergence depends only mildly on these parameters and we give a partial analysis of this phenomenon. We also show that some expensive subsidiary computations required by the new method can be replaced by inexpensive approximate versions of these tasks based on iteration, with virtually no degradation of performance

ISCR Institute for Scientific Computing Research

July 22, 1998

Model-Based Object Recognition Using Algebraic Functions of Views

George Bebis

University of Nevada, Reno

Abstract

The recognition and classification of objects is a spontaneous, natural activity for many biological vision systems. In contrast, building systems capable of recognizing relevant 3D objects in their environment, with accuracy and robustness, has been a difficult and challenging task in computer vision. Object recognition is difficult is because the appearance of an object can have a large range of variation due to photometric effects, scene clutter, changes in shape (e.g., the object is not rigid) but most importantly, due to viewpoint changes. This results in numerous different images even for the same object.

Accommodating variations due to viewpoint changes is a central problem in the design of any object recognition system, and one whose solution is likely to have implications throughout the system. In this talk, we will present a new technique for object recognition based on the recently proposed theory of "algebraic functions of views," which provides a powerful mathematical foundation for tackling variations in the appearance of a 3D object's shape due to viewpoint changes.

July 17, 1998

Synthesis of Linear Phase Multirate Filter Banks for Signal and Image Coding Applications

Chris Brislawn

Los Alamos National Laboratory
brislawn@lanl.gov

Abstract

Wavelet transforms and their digital signal processing counterparts, multirate filter banks, have proven superior to block transform coding for a variety of digital data coding and are now being incorporated into a number of commercial and international standards for audio, still and moving picture data compression. The speaker was involved in drafting the first such wavelet-based standard, the FBI's national specification for coding digitized fingerprint images.

This talk will review the definitions of wavelet transforms and multirate filter banks and indicate how they are used in digital signal and image coding. We will then present a new cascade-form architecture for synthesizing linear phase filter banks of one of the classes commonly used in image compression applications. The new architecture, which improves on earlier work of Vaidyanathan and Nguyen, achieves a reduction of 33% to 50% (asymptotically) in the number of multiplications per unit input needed to implement the filter bank when compared to direct form implementation.

ISCR Institute for Scientific Computing Research

July 10, 1998

Legion – An Applications Perspective

Andrew Grimshaw

University of Virginia
grimshaw@cs.virginia.edu

Abstract

Legion is a reflective metasystem project at the University of Virginia. It is designed to provide users with a transparent interface to resources in a wide-area system, both at the programming interface level and at the user level. Legion addresses issues such as parallelism, fault-tolerance, security, autonomy, heterogeneity, resource management, and access transparency in a multi-language environment. While fully supporting existing codes written in MPI and PVM, Legion provides features and services that allow users to take advantage of much larger, more complex resource pools.

With Legion, for example, a user can run a computation on a supercomputer at a national center while dynamically visualizing the results on a local machine. As another example, Legion makes it trivial to schedule and run a large parameter space study on several workstation farms simultaneously. Legion permits computational scientists to use cycles wherever they are, allowing larger jobs to run in shorter times through higher degrees of parallelization.

July 1, 1998

Hierarchical Approaches for Representing and Visualizing Massive Scientific Data Sets

Bernd Hamann

University of California Davis
Department of Computer Science

Co-Director of the Center for Image Processing and Integrated Computing (CIPIC)
hamann@cs.ucdavis.edu

Abstract

One of the most challenging and important problems that the science and engineering communities are facing today—and even more so in the future—is representing, visualizing, and interpreting very large data sets. Such data sets commonly result from computer simulations of complex physical phenomena (e.g., computational physics, climate modeling, or ocean modeling) and from high-resolution imaging (e.g., satellite imaging or medical imaging).

The technology currently being used to represent massive data sets is inappropriate for interactive and efficient data analysis and visualization. It is impossible for a user of a visualization system to "navigate" through a data set consisting of several million (or billion) data points, and analyze the data set entirely. In this talk, I will present various ideas that seem promising in the context of overcoming some of the problems associated with the visualization of very large data sets. I will point out the necessity of bringing together ideas from approximation theory and geometric modeling (splines), computational geometry (tessellations), optimization (simulated annealing), and other related fields. I will point out avenues for representing massive data sets using hierarchical approaches that facilitate visualization and analysis.

University of California
Lawrence Livermore National Laboratory

ISCR Institute for Scientific Computing Research

June 29, 1998

Least Squares Finite Element Approximations Based on Minus One Inner Product, and Their Analysis

Raytcho Lazarov

Texas A & M University
lazarov@math.tamu.edu

A presentation of joint work conducted with J. Bramble and J. Pasciak

Abstract

The merits and the deficiencies of various least squares formulations for solving boundary value problems will be discussed as a short introduction. Major advantages of the least squares methods are the stability, symmetry and positive definiteness of the corresponding discretizations.

The main objective of this talk is to introduce and to analyze least squares formulations for second order elliptic problems based on a discrete minus one inner product. The proposed formulation does not require the classical condition of Ladyzhenskaya–Babuska–Brezzi. Using these functionals, we develop finite element approximations for non-symmetric and indefinite second order elliptic problems, and for the equations of incompressible and almost incompressible elastic media. The method leads to symmetric and positive definite algebraic problems of optimal rate of convergence.

This approach is based on the recent theoretical advances and practical implementations of multilevel and multigrid methods, multilevel splittings of Sobolev norms, and domain decomposition methods.

CASC Center for Applied Scientific Computing

June 25, 1998

Efficient Elliptic Solvers for Supercomputers Based on High-Performance Microprocessors

Ulrich Ruede

Institut fuer Mathematik,
Universitaet Ausgburg, D-86135 Ausgburg, Germany
Ulrich. Ruede@Math.Uni-Ausgburg.DE

Abstract

We recognize the trend to build even supercomputers using standard microprocessors. Such processors are amazingly fast in terms of computational power. The performance of such systems, however, is often limited by the speed of the memory system. This is a severe bottleneck for all conventionally programmed iterative methods, since they exhibit only limited re-use of data from faster cache memory.

Computational experiments can show the performance of different data structures and algorithmic patterns for current microprocessors and their memory systems. To increase the efficiency, the design of iterative methods must avoid patterns that limit on-chip parallelism, and the algorithms must be restructured such that the re-use of cached data is improved. With a carefully designed multigrid method, a system with one million unknowns can be solved in approximately one second on a single CPU node.

ISCR Institute for Scientific Computing Research

June 19, 1998

High-Performance Thread Libraries: Past and Current Work

Dirk Grunwald

University of Colorado Boulder
grunwald@foobar.cs.colorado.edu

Abstract

Computer runtime systems use thread libraries to provide a way for programmers to express concurrency. On sequential processors, a thread library can be used to mask communication latency by switching threads on communication stalls. On shared memory parallel processors, threads are used to provide work for different processors. In this talk, I'll describe two studies to improve the performance of thread systems.

The first study[1] improves the mechanics of thread libraries, particularly for fine-grained threads. For conventional languages, threads require space for the per-thread stack and must save registers across context-switch boundaries. We used whole-program optimization to reduce the needed space, and, as a side effect, the number of TLB misses encountered by threaded programs. We also used link-time inter-procedural live-register analysis to determine and reduce the number of live registers that need to be preserved around a voluntary context switch. Both of these mechanisms were implemented on a DEC Alpha system using the "OM" binary optimization system from DEC Western Research Labs, and speedups of 15–35% were demonstrated.

The second study[2] was designed to reduce the execution cost and simplify the scheduling of scientific applications on parallel systems. We designed a new thread abstraction, called Iterates, that represents the execution of a subspace of an iteration space. Iterates are linked by dependence constraints that are resolved at execution time. This combination effectively turns conventional applications into “large-grain data flow” execution models, allowing the scheduling system to implement load balancing, affinity scheduling and automatic wave fronting. A prototype system was implemented on a DEC Alpha system, and achieved speedups of 10–40% on various application kernels.

In closing, I'll also mention an on-going project to build a high-performance distributed object system based on the CORBA distributed object model.

[1]Joint work with Richard Neves, now at the IBM T.J. Watson Research Center, Yorktown Heights, NY.
[2]Joint work with Suvas Vajracharya, now at Los Alamos National Laboratory.

June 18, 1998

Reduced-Order Modeling for Muti-Input Multi-Output Linear Dynamical Systems

Roland Freund

Bell Laboratories
freund@research.bell-labs.com

Abstract

Multi-input multi-output time-invariant linear dynamical systems arise in important applications, for example, in the simulation of electronic circuits. The state-space dimensions of such systems can be so large that time-domain integration of the original system would be inefficient or even prohibitive. Instead, the original system is replaced by a suitable reduced-order model whose state-space dimension is small enough that it can be solved numerically.

In this talk, we discuss reduced-order modeling techniques based on matrix-Pade approximation of the frequency-domain transfer function of the multi-input multi-output linear dynamical system. We show how such reduced-order models can be computed in a stable and efficient way via a Lanczos-type method for multiple starting vectors. It is desirable and often crucial that reduced-order models inherit the essential properties, such as stability or passivity, of the original system. Next, we discuss stability and passivity of Pade-based reduced-order models. Finally, we present results of numerical experiments with linear dynamical systems arising in circuit simulation.

ISCR Institute for Scientific Computing Research

June 16, 1998

High-Performance Computational Grids

Ian Foster

Argonne National Laboratory
and University of Chicago

Abstract

A computational grid, like its namesake the electric power grid, provides quasi-ubiquitous access to capabilities that cannot easily be replicated at network endpoints. In the case of a high-performance grid, these capabilities include both high-performance devices (networks, computers, storage devices, visualization devices, etc.) and unique services that depend on these devices, such as smart instruments, collaborative design spaces, and metacomputations.

In this talk, I discuss some of the technical challenges that arise when we attempt to build such grids: in particular, the frequent need to meet stringent end-to-end performance requirements despite a lack of global knowledge or control. I then introduce the Globus Project, a multi-institutional effort that is developing key grid infrastructure components, for authentication, resource location/allocation, process management, communication, and data management. I conclude with a discussion of our experiences developing test beds and applications based on Globus components.

June 12, 1998

Database Middleware for Heterogeneous Data Sources

Laura Myers Haas

IBM Almaden, Garlic Project
LAURA@almaden.ibm.com

Abstract

Businesses today rely on large collections of data stored in diverse systems with differing capabilities. Database middleware systems can provide an integrated view of data stored on various relational systems, as well as a few non-relational sources. They do not help, however, when the data sources to be integrated include a broader range of sources, such as CAD/CAM systems, text search engines, molecular structure databases, or customer-specific repositories.

We describe Garlic, an IBM prototype that allows integration of diverse sources such as the above, and allows new sources to be easily added to an existing installation. Garlic offers the ability to interrelate data from multiple sources with a broad range of querying capabilities, in a single, cross-source query. A significant focus of the project is the provision of support for data sources that provide type-specific indexing and query capabilities, such as text search, or search by molecular structure. Garlic's "wrapper architecture" encapsulates data sources, allowing new sources to be added quickly, and accommodating a wide variety of sources, with a broad range of traditional and nontraditional query processing capabilities. Wrappers model legacy data as objects, participate in query planning, and provide standard interfaces for method invocation and query execution. Garlic provides an object-oriented query language (along the lines of SQL3), and extends database optimization technology to create efficient plans for queries over multiple sources, whatever their query capabilities, using wrapper input.

A Garlic prototype has been operational since 1995, and has been used as the basis of a customer joint study in the pharmaceutical domain. We are currently using Garlic as the basis of a new venture into scientific information management, initially for the petroleum industry. In this talk, we will describe the overall Garlic architecture, as well as our experiences to date when using Garlic.

University of California
Lawrence Livermore National Laboratory

ISCR Institute for Scientific Computing Research

June 11, 1998

The Development of an Agglomeration Multigrid Method for Highly Anisotropic Unstructured Mesh CFD Problems

Dimitri Mavriplis

Research Fellow
ICASE, NASA Langley Research Center
dimitri@icase.edu

Abstract

Agglomeration multigrid methods represent a variant of algebraic multigrid methods that can be applied directly to non-linear problems. This enables the solution of computational fluid dynamics (CFD) problems without the explicit storage of a Jacobian, which leads to large memory savings over the traditional approach of applying algebraic multigrid to the linearized system of equations. Similar to algebraic multigrid methods, agglomeration multigrid methods consist of a setup phase (coarsening) and a solution phase.

For inviscid flow problems where the underlying grid is isotropic, coarsening algorithms based on unweighted graphs have been shown to result in good overall convergence rates. For high-Reynolds number viscous flows, where extreme grid stretching is required to resolve the thin boundary layer and wake regions, directional coarsening is required, which can be implemented using a weighted graph algorithm. However, directional coarsening results in higher complexity coarse level meshes, which in turn requires additional memory and CPU time within a multigrid cycle. To overcome this problem, aggressive coarsening strategies have been developed, where coarse levels of prescribed complexity reduction are generated. To maintain effective convergence rates, this strategy is coupled with a locally implicit multigrid smoother, which operates on locally constructed lines in the unstructured mesh, using a weighted graph algorithm.

Results will be presented that demonstrate fast convergence rates for aerodynamic Navier-Stokes flows that are independent of the grid stretching. In addition, the parallelization of this solver using the MPI message-passing library will be described, and scalability results on the CRAY T3E and SGI Origin 2000 will be given. Finally, the solution of a large-scale problem involving over 13 million grid points on the T3E will be described.

May 29,1998

Three Modest Proposals in Aid of Scaling PDE Solvers to Teraflops: Asynchronicity, Tensoricity, and Memtropy

David E. Keyes

Old Dominion University
Computer Science Department, and ICASE, NASA
Langley Research Center
keyes@cs.odu.edu

Abstract

Asynchronous iterative methods have somewhat of a "bad name" in the theoretical parallel algorithms literature, where worst case estimates for linear problems are pessimistic. Asynchronous methods, however, may deserve renewed consideration for heterogeneous nonlinear problems at ASCI scales. Heterogeneities (e.g., shocks, flame fronts, plastic zones) lead to "nonlinear stiffness," (widely varying magnitudes in the nonlinearity tensor, the term after the Jacobian in the multivariate Taylor expansion upon which Newton's method is based).

Many important problems are strongly nonlinear in a small region that is embedded in an ambient region that is only weakly nonlinear. Treating heterogeneous problems bulk synchronously, as with the speaker's favorite Newton-Krylov-Schwarz methods, at high processor granularities either wastes cycles in the near-linear regions or enslaves load balancing to the evolving physics. Alternatively, it should be relatively easy to accommodate asynchronous forms of nonlinear Schwarz iteration through multithreading, the natural software model for interactive visualization and computational steering. Each process "publishes" updates of solution values for on-demand inspection by processes governing neighboring subdomains, not by a global hand-shaking exchange.

Exploiting the possibilities of nonlinear Schwarz in an automated way requires a pointwise metric, analogous to a discretization error estimate in an adaptive discretization code, which quantifies the degree of nonlinearity. We propose a metric, "tensoricity," for this purpose. Orthogonally, but motivated by the same challenge of developing PDE solvers effective at Teraflops scales, we propose a new metric, "memtropy," to provide architecture-independent rankings of data locality in programs destined for deep-memory hierarchy machines. No architecture-independent metric can be expected to be a reliable performance predictor on a given machine, but memtropy should be convenient in assessing trends when laying out data for non-flat memory systems, generally.

No performance data justifying the approaches advocated herein will be given yet; however, the necessity of moving beyond synchronous SPMD implementations of PDE solvers will be argued from performance data from large-scale aerodynamics runs. Moreover, the advocated looser approaches potentially lead to better machine utilization in the multi-physics computations ultimately required in ASCI.

ISCR Institute for Scientific Computing Research

May 22, 1998

A Subspace-Based Model for Information Retrieval with Applications in Latent Semantic Indexing

Hongyhan Zha

Pennsylvania State University

Abstract

A theoretical foundation for latent semantic indexing (LSI) is proposed by adapting a model first used in array signal processing to the context of information retrieval using the concept of subspaces. It is shown that this subspace-based model, when coupled with the minimal description length (MDL) principle, leads to a statistical test to determine the dimensions of the latent-concept subspaces in LSI.

The effect of weighting on the choice of the optimal dimensions of latent-concept subspaces is illustrated. It is also shown that the model imposes a so-called "low-rank-plus-shift structure" that is approximately satisfied by the cross-product of the term-document matrices. This structure can be exploited to give a more accurate updating scheme for LSI, and to correct some of the misconception about the achievable retrieval accuracy in LSI updating. It is further demonstrated that based on the low-rank-plus-shift structure a divide-and-conquer method can be devised to compute the partial singular value decomposition (SVD) of a large sparse term-document matrix. Possible extensions of the model to improve retrieval accuracy are also pointed out.

CASC Center for Applied Scientific Computing

May 19, 1998

Problem-Solving Environments and NetSolve: A Network Server for Solving Computational Science Problems

Jack Dongarra

University of Tennessee

Abstract

This talk presents a system called NetSolve that allows users to access computational resources, such as hardware and software, distributed across the network. This project has been motivated by the need for an easy-to-use, efficient mechanism for using computational resources remotely. Ease of use is obtained as a result of different interfaces, some of which do not require any programming effort from the user.

Good performance is ensured by a load-balancing policy that enables NetSolve to use the computational resource available as efficiently as possible. NetSolve offers the ability to look for computational resources on a network, choose the best one available, solve a problem (with retry for fault-tolerance) and return the answer to the user.

ISCR Institute for Scientific Computing Research

May 1, 1998

Project VisualEyes: Integrated Data Synthesis, Analysis and Visualization of Parallel Adaptive Simulations

Chandrajit L. Bajaj

University of Texas at Austin

Abstract

Even with the use of parallel supercomputers, conventional approaches to simulation and visualization are weakly suited to terascale problem sizes, where discretized domain models extracted from large images are much too large for interactive analysis, visualization and where physical simulations could take days to weeks. One experiment is to extract and employ progressively encoded, hierarchical meshes to support both adaptive simulation and interrogative visualization.

The VisualEyes project is focused on research in some of the core technologies (e.g., compression, adaptive meshing, interrogative visualization, and error estimation) as well as in integrating parallel simulation, with data analysis and collaborative visualization of multivariate scalar, vector, tensor fields. In this talk I shall present details of our progressive mesh encoding scheme, as well as the use of vector field topology for several applications in interrogative visualization. This project is driven by close cooperation with TICAM application developers in composite materials, electromagnetic scattering, and reservoir modeling.

CASC Center for Applied Scientific Computing

April 28, 1998

Coherent Structure Theory and Voting Rules

Dean H. Judson

University of Nevada, Reno
Statistics and Research Methods Laboratory
and Social Psychology Ph.D. Program

Abstract

This presentation is a case study in systems isomorphy or systems analogy. I present the basic terms of coherent structure theory, which was originally developed for engineering reliability applications. I then show how the elements of coherent structures (components, minimal paths, modules, and monotonic structures) can be given a coalition or voting interpretation (decision-makers, minimal winning coalitions, subcommittees, or irrelevant group members).

I describe some basic theorems derivable from coherent structure axioms, and I discuss some implications of interpreting the theorems with respect to voting rules. Finally, I describe a recent project involving inferring the underlying coherent structure from a sample of data on components and system outcomes.

ISCR Institute for Scientific Computing Research

April 20 and 27, May 4 and 11, 1998

Large-Scale Simulations in Fortran 95: An Object-Based Approach

Paul F. Dubois

Lawrence Livermore National Laboratory
X-Division
dubois1@llnl.gov

Abstract

The course is intended for experienced Fortran 77 users. After this "free upgrade" to your skills, you should be able to use the full power of Fortran 95 and understand how to use an object-based approach to large-code architecture.

If you do not currently have any Fortran 90/95 reference materials, consider the following: *Fortran 90 Programming*, T.M.R. Ellis, Ivor R. Philips, and Thomas M. Lahey, Addison-Wesley, Reading, MA., 1994, ISBN 0-201-054446-6. If you search an on-line bookstore such as amazon.com with the word "Fortran" you'll find plenty more to choose from. If you have access to a Windows 95/NT machine, strongly consider the Digital Visual Fortran integrated environment. The book that comes with this is very good. (If you get one; one person reports that he didn't get anything except a CD.)

April 17, 1998

Representation and Image Comparison Metrics for Large Data Sets

Raghu Machiraju

Mississippi State University
NSF Engineering Research Center on Computational Field Simulation

Abstract

In this talk I address two important operations needed for the visualization of large (terascale?) data sets, namely compressed domain representation, and metrics for comparing data sets and images. We first show how multiresolutional schemes based on wavelets can be used to represent large data sets.

Wavelets are used to detect coherent structures, which persist across scales. The sub-combining technique can be employed to identify regions in a volume that contain significant structures. These regions can be spatially partitioned (with octree) into blocks and each block can be coded again with a wavelet transform. The coefficients can now packed into a bitstream in such a way that a ranking of perceptually significant structures is obtained. We show an example use of this representation for rendering iso-surfaces.

Using the same multiresolution representation scheme, we design image comparison metrics. We describe how the metrics respond to operators normally used in rendering. Finally, we show how these metrics can be extended to multi-part metrics that can be used to compare images and data sets alike and provide a better understanding of differences. Such metrics find use in navigating and browsing

ISCR Institute for Scientific Computing Research

April 17, 1998

Change Detection in Data Warehousing and a Performance Study

Nabil Rashad Adam

and

Igg Adiwijaya

Center for Information Management Integration and Connectivity (CIMIC),
Rutgers University, Newark, NJ

Abstract

The change detection problem is encountered in various areas such as digital libraries and electronic commerce, specifically in data warehousing and Internet Web pages, among others. In such systems, the change detection problem manifests itself into two possible forms. First, there is a need to detect and propagate relevant changes to the underlying information sources to ensure data consistency. Second, such systems encounter continuous change not only in the content of Web pages but also to the Web pages themselves as new pages may be added and existing pages may be removed.

In this talk we address these two forms of change detection, i.e., detecting and propagating changes to relevant data (information sources in the former case and Web pages in the latter case).

CASC Center for Applied Scientific Computing

April 14, 1998

Multiresolution Algorithms in Computer Graphics

Peter Schroeder

California Institute of Technology

Abstract

Computational techniques based on wavelets and more general multiresolution approaches have made tremendous inroads into computer graphics applications. Examples from modeling, simulation, and rendering include hierarchical surface editing, illumination computations, and level-of-detail display, among others. The success of these techniques is based on the favorable scaling properties of multiresolution transforms, their flexible time/accuracy tradeoffs, and recent advances in very general multiresolution construction techniques.

In this talk I will give an overview of the ideas behind these constructions, and use application examples from radiative transport, large -scale data compression, and arbitrary topology surface modeling, to illustrate their use effectiveness.

ISCR Institute for Scientific Computing Research

April 10, 1998

Robust Simplification Methods for Triangle and Tetrahedral Meshes

Kenneth I. Joy

University of California, Davis
Computing Science Department Center for Image Processing and Integrated Computing

Abstract

One of the most critical and fundamental research problems encountered in the analysis and visualization of massive data sets is the development of methods for storing, approximating, and rendering large sets of data efficiently. The problem is to develop different representations of the data set, each of which can be substituted for the complete set depending on the requirements of the analysis or the visualization technique.

In this talk we present new methods for the construction of multiple levels of triangle and tetrahedral meshes. Starting with an initial, high-resolution triangulation, we construct coarser representation levels by collapsing simplices in the mesh. Based on weights defined for each simplex, we identify the simplex whose elimination would cause a minimal increase in error, and collapse this object. Weights are stored for individual simplices and are updated as the mesh is simplified. Different strategies can be used in the simplification process to produce hierarchies of meshes that can be used for different applications. These methods result in a hierarchical data description suited for the efficient visualization of large data sets at varying levels of detail.

CASC Center for Applied Scientific Computing

March 30, 1998

The Challenge of Fermion Monte Carlo

Malvin H. Kalos

Cornell University
Theory & Simulation Science & Engineering Center

Abstract

We will review the fundamental challenge of fermion Monte Carlo for continuous systems, the "sign problem," and some of the proposals that have been made for its solution. In particular, we describe methods that depend upon the use of correlated dynamics for ensembles of correlated walkers that carry opposite signs. We discuss the algorithmic symmetry between such walkers that must be broken to create a method that is both exact and as effective as for bosonic systems. We explain the concept of marginally correct dynamics. Stable overlaps with an antisymmetric trial function given by such dynamics correspond to the correct fermion ground state.

Many-body harmonic oscillator problems are particularly tractable: Their stochastic dynamics permit the use of regular geometric structures for the ensembles. These structures are stable when appropriate correlations are introduced, and avoid the decay of signal-to-noise that is a normal characteristic of the sign problem.

Finally, we outline a generalization of the method for arbitrary potentials, and describe the progress in treating few electron systems.

University of California
Lawrence Livermore National Laboratory

ISCR Institute for Scientific Computing Research

March 27, 1998

Computational Analysis in Orthopedic Design

Ann Hollister

Louisiana State University Medical Center

and

Louise Focht

Avanta Orthopedics, Inc.

Abstract

Anne Hollister, orthopedic surgeon at Louisiana State Medical Center, and Louise Focht, co-founder of Avanta Orthopedics, Inc. (San Diego) will present clinical and industry perspectives on the applications of computational modeling in prosthetic implant design for human joint replacements. Traditional design processes involve CAD development but do not include extensive computational analysis. One result has been that design flaws often do not appear until after many prostheses have been implanted, and patients have suffered from premature failure of these joint replacements. Properly executed, computational analysis can predict likelihood of failure of joint replacements and, when incorporated early in the design process, can improve the design quality.

Hollister is a practicing surgeon and has designed several prosthetic joints. Avanta Orthopedics is a company specializing in replacements for small joints, such as those found in the hand, elbow, and foot.

March 20, 1998

Multigrid Methods for Unstructured Grids and for Convection–Diffusion

Jinchao Xu

Penn State University
Mathematics Department
xu@math.psu.edu

Abstract

After a general introduction of multigrid methodology, a number of techniques will be presented in this talk on how a finite element equation posed on an unstructured grid in two or three dimensions can be solved within optimal computational complexity by special multigrid methods such as nonnested multigrid methods, auxiliary space methods and agglomeration methods. If time allows, some new discretization and iterative techniques will be discussed on convection dominated problems. Both theoretical and numerical results will be reported.

University of California
Lawrence Livermore National Laboratory

ISCR Institute for Scientific Computing Research

March 19, 1998

Solving Eigenvalues Problems

Jesse L. Barlow

Pennsylvania State University
Department of Computer Science and Engineering
barlow@cse.psu.edu

Abstract

This talk gives a discussion of recent work in the solution of eigenvalue problems. Two recent classes of results are surveyed: accuracy and perturbation theory, and methods for solving large-scale eigenvalue problems. In the first area, we discuss new ways of characterizing accurate solution of the symmetric eigenvalue and singular value problems. Traditionally, the singular value problem has been viewed through the window the symmetric eigenvalue problem, but it turns that it is better to characterize the accurate solution of the symmetric eigenvalue problem through the window of the singular value decomposition. In the solution of large--scale eigenvalue problems, we look at some recent results on approximations generated by Lanczos methods.

March 9, 1998

Revolutionizing Parallel Programming through Better Hardware

John Feo

Tera Computer Company

Abstract

The commercial availability of MPP and clustered SMP systems has made very high peak performance systems readily available to scientists. However, the difficulty of programming for data locality and the limitations of medium- and coarse-grain parallelisms have limited the usefulness of these systems. The Tera MTA is a scalable, shared memory, MIMD computer that supports very fine-grain synchronization. It has no local memory and no caches. Each processor is 128 virtual processors feeding a single instruction pipeline. As long as one of the virtual processors can issue an instruction every cycle, the processor remains fully utilized. Thus, parallelism is the only limiting factor on the MTA. Data locality, granularity, and task scheduling for load balancing are non-issues. The MTA sustains high execution rates over a wide spectrum of applications, while reducing programming costs and encouraging parallel algorithm development.

In this talk, I will describe the overall architecture of the MTA and the current status of the SDSC machine. I will explain how one develops and optimizes code for the system using the ASCI benchmark Sweep3D and other applications as time permits. I will give some preliminary performance numbers that compare quite favorably with other HPC systems.

ISCR Institute for Scientific Computing Research

February 26, 1998

Information Retrieval via Limited-Memory Matrix Methods

Tamara G. Kolda

University of Tennessee
Department of Computer Science and Oak Ridge National Laboratory (ORNL) Householder Fellow, Mathematical Sciences Section
kolda@msr.epm.ornl.gov

Abstract

With ever-larger collections of documents available electronically, a need has arisen for fast and efficient search engines. Latent semantic indexing (LSI) approximates a matrix representing a document collection using the truncated singular-value decomposition (SVD); this allows automatic recognition of latent relationships between words and leads to a more efficient search engine. We propose replacing the SVD with what we call the semi-discrete decomposition (SDD). The resulting SDD-based LSI performs as well as the SVD-based method, requires substantially less storage, and processes queries faster. Furthermore, the SDD is easy to update when new documents are added to the collection.

CASC Center for Applied Scientific Computing

February 24, 1998

An Adaptive Projection Method for Modeling Low Mach Number Flows

John B. Bell

Lawrence Berkeley National Laboratory (LBNL)
Center for Computational Science and Engineering (CCSE)
JBBell@lbl.gov,

Abstract

We describe an adaptive projection algorithm for low Mach number flows. The basic approach uses hierarchical grids that are refined in both space and time. In this presentation we will review the single grid projection methodology and discuss the key issues in the algorithm that must be addressed in developing an adaptive algorithm. With these issues in mind, we discuss the design principles that we use to solve partial differential equations on adaptive grids.

We then describe how these principles are used to develop an adaptive version of the projection algorithm. Numerical examples illustrating the capabilities of the method will be presented.

ISCR Institute for Scientific Computing Research

February 20, 1998

Multi-Source Data Analysis in Science and Engineering

Samuel P. Uselton

NASA Ames
MRJ Technology Solutions
uselton@nas.nasa.gov

Abstract

As digital data acquisition becomes easier, cheaper and more pervasive, and computational simulations gain increasing fidelity and detail, many activities can benefit from the combined analysis of data from several sources. Building a useful multi-source analysis system requires solving many problems, some of a pragmatic, engineering nature, and some of a more basic nature. The issues seem remarkably similar whether the application is weather modeling, environmental assessment and remediation planning, oil exploration and production, or engineering design processes. The problems to be addressed range from efficient access to large amounts of data from multiple heterogeneous sources, to design of user interfaces, and inventing visualization techniques. Integrated display and comparative analysis of relevant data is interesting and relatively unexplored.

Our multi-source visualization (MSV) project uses a specific problem domain, concurrent design of aircraft, to focus research and development efforts in this area. Frequent contacts with the Earth observation system (EOS) project, meso-scale atmospheric modeling researchers, and nanotechnology researchers, among others, keep broader needs in view.

All the work done by the Data Analysis Group of the NAS Division at NASA Ames Research Center is relevant and useful to this project. This group's work ranges from specific visualization techniques, through innovative user interfaces, the software engineering required to build complete systems, systems level performance improvements, effective access to very large data sets, and exploitation of large heterogeneous collections of scientific and engineering data. Much of this work will be touched upon, but the focus will be on the work that is specifically driven by the need for a variety of users to exploit common collections of data from many diverse sources.

This work is very much in progress. There are results to show, but also problems not yet satisfactorily resolved. This talk will be profusely illustrated, of course.

CASC Center for Applied Scientific Computing

February 20, 1998

Simulating Turbulence and Turbulent Convection on SMP Clusters, with the PPM Gas Dynamics Code

Paul Woodward

University of Minnesota
Laboratory for Computational Science and Engineering (LCSE)

Abstract

The PPM gas dynamics code is being used in an extensive program of simulations of turbulent thermal convection in stars. During the last year it has been possible to treat entire 3D model stars, using this code on large SMP clusters. Using the ASCI machines at Los Alamos, detailed simulations of the compressible turbulence in its own right have also been carried out using the same techniques on grids of up to a billion cells. This work, the product of an interdisciplinary team at the University of Minnesota's Laboratory for Computational Science & Engineering collaborating with both Livermore and Los Alamos, gives us the opportunity to look inside stars and to observe the complex interactions of thermal convection with both rotation and pulsation of the stars.

The simulations of turbulence also offer unique views at unprecedented resolution of the dynamical processes at work. The methods used to make these calculations efficient on SMP clusters will be discussed and early visualizations of the results presented. These large simulations produce multi-terabyte data sets. Techniques used to deal with these data volumes will be discussed and directions of further research presented.

University of California
Lawrence Livermore National Laboratory

ISCR Institute for Scientific Computing Research

February 11, 1998

Tecolote: An Object-Oriented Framework for Hydrodynamics

Kathy Holian

Los Alamos National Laboratory
Software Design and Development
ksh@lanl.gov

Abstract

Tecolote is an object-oriented framework for both developing and accessing a variety of hydrodynamics models. It is written in C++, and is in turn built on the Parallel Object-Oriented Methods and Applications (POOMA) framework. The Tecolote framework is meant to provide modules (or building blocks) to put together hydrodynamics applications that can encompass a wide variety of physics models, numerical solution options, and underlying data storage schemes, although with only the necessary modules activated at runtime . Tecolote has been designed to separate physics from computer science, as much as humanly possible.

The POOMA framework provides fields in C++ to Tecolote that are analogous to Fortran 90-like arrays in the way that they are used, but that, in addition, have underlying load balancing, message passing, and a special scheme for compact data storage. The POOMA fields can also have unique meshes associated with them that can allow more options than just the normal, regularly spaced Cartesian mesh. They also permit one, two, or three dimensions to be immediately accessible to the code developer and code user.

Having a framework in which to develop hydrodynamics methods furnishes numerous advantages. The most obvious is that one can build upon models (both physical and numerical) that have already been developed and tested. It is also advantageous for the physics programmer to have complicated computer science issues already solved, yet abstracted away. The types of computer problems that need to be solved these days, but that the physicist does not want to have to deal with, include such things as message passing and load balancing.

We have endeavored wherever possible to write the framework in as object-oriented a manner as possible so that model developers can enjoy maximal code reuse. This has also led to a nearly complete separation of the computer science from the physics coding. The framework has also been designed to be completely portable across a wide variety of platforms.

I will discuss the philosophy of the Tecolote framework, and will present a summary of the overall design, including a discussion of some of the objects that were used to put the framework together. The first hydrodynamics option using the framework that we are developing is a multi-material Eulerian code. This will allow us to compare with an existing hydrocode written in Fortran 90. I will present comparisons of both timings for large problems, and of possible maximum size of problem on the computer platforms of interest at LANL.

February 6, 1998

Parallel Adaptive Mesh Refinement

Dan Quinlan

Los Alamos National Laboratory
Scientific Computing Group
dquinlan@lanl.gov

Abstract

Adaptive mesh refinement is a numerical technique for locally tailoring the resolution of computational grids. AMR permits the addition of finer grids to the global computational grid in an adaptive way so as to permit locally more accurate computations or the removal of global error introduced by local singularities. AMR as a numerical technique, is largely independent of the equations being solved, though numerous numerical and algorithmic issues are involved and are the subject of significant research. Unfortunately, AMR is not common place due largely to its inherent complexity.

Adaptive mesh refinement computations are complicated by their dynamic nature. In the serial environment, they require substantial infrastructures to support the regridding processes, intergrid operations, and local bookkeeping of positions of grids relative to one another. In the parallel environment, the dynamic behavior is more problematic because it requires dynamic distribution support and load balancing. Parallel AMR is further complicated by the substantial task parallelism, in addition to the obvious data parallelism, this task parallelism requires additional infrastructure to support efficiently. The degree of parallelism is typically dependent upon the algorithms in use and the equations being solved. Different algorithms have significant compromises between computation and communication. Substantial research work is often required to define efficient methods and suitable infrastructure. The purpose of this talk is to introduce AMR++ as an object-oriented library which forms a part of the Overture framework, a much larger object-oriented numerical framework developed and supported at Los Alamos National Laboratory and distributed on the Web for the last several years.

The parallel issues in AMR are the subject of special attention within AMR++. Overture provides simple mechanisms for the specification of the distribution of Overture grid data (via P++ distribution mechanisms). These distribution mechanisms are used by AMR++ and the multilevel load balancer (MLB) to define the distribution of the adaptive grid. Different AMR algorithms place different requirements upon the distribution of the adaptive grid in the parallel environment. AMR++ supports a couple of different distribution mechanisms specialized for the most common sorts of adaptive mesh refinement solution methods. The underlying mechanisms for defining new distributions, however, are readily available to the user as well. This allows for the tailoring of AMR++ parallel support for different sorts of applications or for research on distribution mechanisms more generally.

ISCR Institute for Scientific Computing Research

January 16, 1998

Adaptive Finite Element Methods on Manifolds, with Applications in Elasticity and Relativity

Michael Holst

University of California Irvine
Department of Mathematics
mholst@uci.edu

Abstract

We discuss the numerical treatment of coupled nonlinear elliptic systems on manifolds. Such systems arise for example, in elasticity models of biological membranes, and in general relativistic models of massive objects. We begin by reviewing some differential geometry, and by taking a brief look at these two applications. Weak formulations of covariant nonlinear elliptic systems on manifolds are then examined.

Finite element approximation theory on manifolds is then discussed, and a computer implementation called MC is described. MC is a dimension-independent, simplex-based, ANSI-C finite element code for the numerical treatment of covariant differential operators on d-manifolds (d = 2,3,...). MC also implements several of the features in the popular package PLTMG (designed for 2D problems in the plane), including a posteriori error estimation, adaptive simplex subdivision, global Newton methods, continuation, and multilevel methods. We describe some of the details of MC, and present some numerical examples for an elasticity problem, and for the elliptic constraints in the Einstein equations.

CASC Center for Applied Scientific Computing

December 19, 1997

Death and Taxes, Nets and Caches: Facing Inevitabilities in Parallel CFD Simulation

David E. Keyes

Old Dominion University
ICASE, NASA Langley Research Center
keyes@cs.odu.edu

Abstract

Demands for massive memory and high speed typically accompany one another in scientific and engineering computations, linking space to time in algorithm design. Some degree of programmer control must be exerted over data layout in coding for scalable distributed memory machines (even when the memory is accessible through the model of a global shared address space).

Fortunately, the laws of nature often cooperate with a basic scaling law of computer architecture: The magnitude of interaction between two degrees of freedom in a physical system decays with their spatial separation; therefore, the frequency and volume of data exchange between different points in the computational domain can be allowed to decay with distance in a trade-off involving memory access overhead and the precision required in a final result (or the rate of convergence required from a preconditioner).

For model problems, this trade-off has been formalized in convergence theorems. We have been exploring it primarily experimentally, applying domain decomposition preconditioners to structured-grid and unstructured-grid problems in computational aerodynamics and acoustics, and keeping Amdahl's Law at bay through the benefits of cache locality. (See http://www.cs.odu.edu/~keyes/nsf/.)

ISCR Institute for Scientific Computing Research

December 16, 1997

Challenges of Future High-End Computing

David H. Bailey

NASA Ames Research Center

Abstract

The next major milestone in high-performance computing is a sustained rate of one Pflop/s (also written one petaflops, or 1015 floating-point operations per second). In addition to prodigiously high computational performance, such systems must of necessity feature very large main memories, as well as comparably high I/O bandwidth, and huge mass storage facilities. The current consensus of scientists who have studied these issues is that "affordable" petaflops systems may be feasible by the year 2010, assuming that certain key technologies continue to progress at current rates.

One important question is whether applications can be structured to perform efficiently on such systems, which are expected to incorporate many thousands of processors and deeply hierarchical memory systems. To answer these questions, advanced performance modeling techniques, including simulation of future architectures and applications, may be required. It may also be necessary to formulate "latency tolerant algorithms" and other completely new algorithmic approaches for certain applications. This talk will give an overview of these challenges.

CASC Center for Applied Scientific Computing

December 12, 1997

Dynamic Distributed Arrays: An Infrastructure for Parallel Implementations of Adaptive Computational Algorithm

James Browne

University of Texas at Austin
Department of Computer Science

Abstract

A dynamic distributed array is an data abstraction that may be expanded or contracted in response to the requirements of an adaptive algorithm but which retains the operational semantics of a normal Fortran array even while distributed across multiple processors. This talk will define and describe two implementations of dynamic distributed arrays.

The first, the hierarchical dynamic distributed array or HDDA, is the basis of an infrastructure for parallel implementation of computations based on structured adaptive meshes. The second, scalable dynamic distributed array or SDDA, is the basis of an infrastructure for parallel implementations of computations based on unstructured adaptive meshes.

The HDDA and the SDDA are built on a common conceptual foundation. Use of the SDDA will be illustrated by formulation of an hp-adaptive finite element computation. Results reported are largely the work of Manish Parashar and Carter Edwards.

University of California
Lawrence Livermore National Laboratory

ISCR Institute for Scientific Computing Research

December 9, 1997

Accurate Discretizations and Efficient Solvers for Heterogeneous Groundwater Flow Equations

Thomas F. Russell

University of Colorado, Denver
Department of Mathematics

Abstract

The ability to compute accurate velocities is important for applications of flow and transport codes. In highly heterogeneous porous media, this task is a difficult one for standard numerical methods. Irregular geological features, which suggest the use of irregular grids, and variable directions of anisotropy add to the challenge. Mixed finite element methods, which solve a continuity equation and a Darcy equation simultaneously for shape–function representations of pressure and velocity, can overcome these obstacles and produce accurate results.

The discrete formulations, however, are often complex, and the linear algebraic equations are not amenable to standard solvers and are difficult to solve as efficiently as equations from other methods. This has inhibited practical applications, especially in 3D. A control–volume variant of the lowest-order Raviart–Thomas mixed method is presented for general logically rectangular grids (2D quadrilaterals, 3D hexalaterals). The Darcy equation is enforced on a cell-sized "tank" (control volume) around each degree of freedom of the velocity (edge in 2D, face in 3D). The discrete equations are simple, easy to implement, and involve only cell pressures and edge or face fluxes (integrated normal velocities).

Numerical tests in 2D show second-order convergence of the edge fluxes whenever the exact solution is not singular. The accuracy far exceeds that of commonly used methods. Also presented is an efficient solver for the 3D equations. It uses a convenient basis for the divergence-free velocity shape functions to reduce the equations to a symmetric positive–definite system of smaller size than has been possible previously. This enables the 3D equations to be solved with effort comparable to that for other methods.

CASC Center for Applied Scientific Computing

November 11, 1997

The Hierarchical Basis Multigrid Algorithm and Incomplete LU Decomposition

Randolph E. Bank

University of California, San Diego
Department of Mathematics

Abstract

We present a new family of multigraph algorithms, ILU-MG, based upon an incomplete sparse matrix factorization using a particular ordering and allowing a limited amount of fill-in. While much of the motivation for multigraph comes from multigrid ideas, ILU-MG is not an typical algebraic multilevel method. The graph of the sparse matrix A is recursively coarsened by eliminating vertices using a graph model similar to Gaussian elimination. Incomplete factorizations are obtained by allowing only the fill-in generated by the vertex parents associated with each vertex.

Because they are graph based, in principle multigraph methods can be applied to general sparse matrices, and in particular do not require a grid hierarchy. We will present some numerical examples that show the behavior of the method is similar to the classical HBMG iteration.

University of California
Lawrence Livermore National Laboratory

November 10, 1997

Multigrid and ILU Methods for Large Scale Computing

Sing-Lok (Justin) Wan

University of California, Los Angeles
Department of Mathematics

Abstract

This talk consists of two parts. In the first, we present a robust interpolation for multigrid methods based on the concepts of energy minimization and approximation. The formulation is general; it can be applied to any number of dimensions and to general computational domains. In 1D, we prove that the convergence rate of the resulting multigrid method is independent of the PDE coefficient and of the mesh size. In 2D, we demonstrate numerically the effectiveness of the multigrid method by applying it to rough coefficient problems and an unstructured grid problem.

In the second part, we present a parallelization of the ILU algorithm. The idea is based on multicoloring. We describe a parallel coloring algorithm by Jones and Plassmann and show how it can be applied to constructing incomplete LU factors for ILU(0). We also discuss an extension to ILU(1) and the issue of the memory allocation.

University of California
Lawrence Livermore National Laboratory